Review

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics

Leveraging the urban–rural divide for epigenetic research

Héléne T Cronjé*,[1], Hannah R Elliott[2,3], Cornelie Nienaber-Rousseau[1] & Marlien Pieters[1]

[1]Centre of Excellence for Nutrition, North-West University, Potchefstroom Campus, Potchefstroom, 2520, North-West Province, South Africa

[2]MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN, UK

[3]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK

*Author for correspondence: toinet.cronje@sund.ku.dk

Urbanization coincides with a complex change in environmental exposure and a rapid increase in noncommunicable diseases (NCDs). Epigenetics, including DNA methylation (DNAm), is thought to mediate part of the association between genetic/environmental exposure and NCDs. The urban–rural divide provides a unique opportunity to investigate the effect of the combined presence of multiple forms of environmental exposure on DNAm and the related increase in disease risk. This review evaluates the ability of three epidemiological study designs (migration, income-comparative and urban–rural designs) to investigate the role of DNAm in the association between urbanization and the rise in NCD prevalence. We also discuss the ability of each study design to address the gaps in the current literature, including the complex methylation-mediated risk attributable to the cluster of forms of exposure characterizing urban and rural living, while providing a platform for developing countries to leverage their demographic discrepancies in future research ventures.

First draft submitted: 5 February 2020; Accepted for publication: 27 April 2020; Published online: 13 July 2020

Keywords: DNA methylation • epigenetic epidemiology • epigenetics and disease • LMIC • noncommunicable disease • urbanization

Cancer, Type II diabetes, respiratory disorders and cardiovascular diseases (CVDs) are collectively responsible for 80% of the 41 million deaths caused by noncommunicable diseases (NCDs) each year [1]. These diseases result from the interplay between fixed genetic [2] and modifiable environmental and behavioral factors [3]. It is the modifiable factors, such as pollution and industrial toxin exposure, physical activity, diet, smoking and alcohol consumption, that are the current focus of NCD prevention and education worldwide [4].

Low- and middle-income countries (LMICs) currently carry 85% of the global NCD death toll. Adults in these regions face twice the risk of NCD mortality than their counterparts living in high-income countries (HICs) [1]. Although this discrepancy is partly accounted for by the lack of adequate healthcare and infrastructure, the main driver of the disproportionate burden of NCDs has been ascribed to the globalization of lifestyles and environmental changes resulting from the rapid rate of urbanization experienced by LMICs. At the same time, these regions remain vulnerable to malnutrition and infectious diseases and are therefore considered to carry the 'double burden of disease' [4,5]. Urban areas are broadly defined by their dense population, commercial activity, nonagricultural employment, level of available education and infrastructure [6]. The strong link between the degree of urbanization (urbanicity) and NCDs supports the notion of environmental factors being instrumental in the etiology of NCDs, rather than purely inherited risk [5]. The possibility exists that differences in epigenetic regulation between urban- and rural-dwelling individuals could be the basis by which the differences in environmental exposure shift disease prevalence with urbanicity.

The epigenome, unlike the genome, is environmentally modifiable and involves the alteration of gene expression without altering the genes themselves [7]. It is also influenced by genetic variation [8]. Epigenetic mechanisms include DNA methylation (DNAm), miRNAs, histone and chromatin modifications [7]. DNAm is currently the

Future Medicine

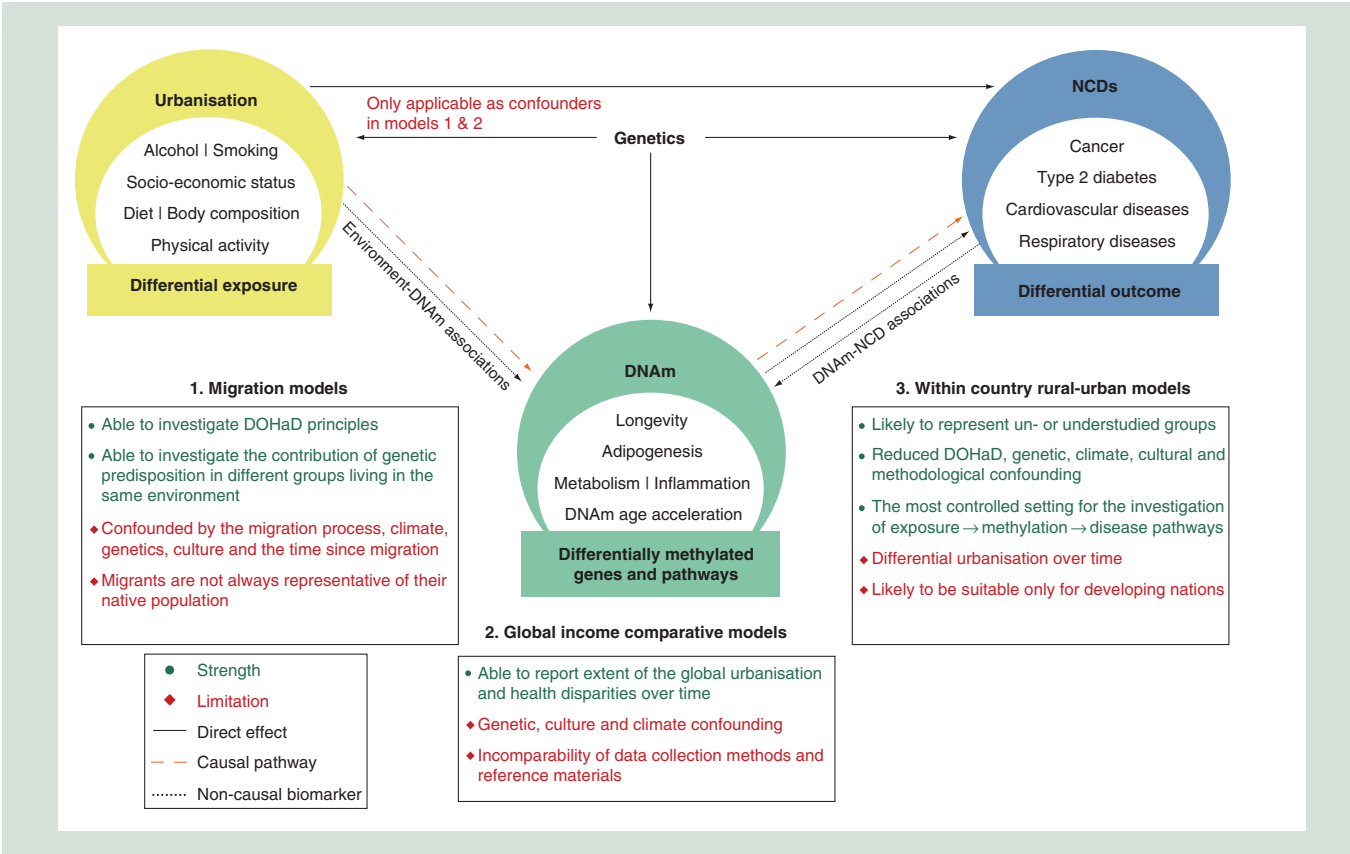

Figure 1. The role of DNA methylation in mediating the association between urbanization and noncommunicable diseases and the strengths and limitations of different study designs aimed at investigating these associations.
DNAm: DNA methylation; DoHaD: Developmental origins of health and disease; NCD: Noncommunicable disease.

most studied epigenetic modification and has been implicated in the etiology and progression of several NCDs over and above genetic predisposition [9]. Methylation differences can be: biomarkers of exposure that do not affect disease, part of the causal pathway between exposures and disease or a biomarker of current disease [9]. Apart from its potential role as a disease mediator, the plasticity of DNAm also makes it a valuable topic of investigation owing to its intervention potential in preventative care [10].

To date, most methylation studies have investigated only single exposures or disease outcomes. This does not take into account that an individual, in any given environment, experiences a combination of exposures simultaneously such as those clustered together in rural and urban landscapes. In this review, we evaluate the ability of three epidemiological study designs to investigate the role of DNAm in the association between urbanization and the NCDs. Investigating urbanization, as a well-defined cluster of exposures, could allow for a better understanding of methylation's role in the global urbanization–NCD trend. First, we review the current evidence for the urbanization–NCD, urbanization–DNAm and DNAm–NCD relationships as the theoretical backdrop to the research models discussed. We then discuss and compare the migration, income-comparative and urban–rural study designs in terms of the questions they are best suited to answer, suitable cohorts and their respective strengths and weaknesses. Figure 1 provides an illustrated summary of this review.

Contextualizing methylation

Genetics and environmental exposures both influence phenotype, partly through epigenetic modifications. Genetic architecture affects DNAm in terms of the availability of cytosine-phosphate-guanine (CpG) sites, the efficacy of methylation-related enzyme expression [8,11], how DNAm responds to exposure [12] and inherent NCD risk [2]. Environmental exposures can be either external (behavioral or lifestyle-related factors such as diet, exercise and air purity) or internal (metabolic and biochemical factors such as inflammation and adiposity). The external

environment's role in methylomic variance is best observed in monozygotic twins that start their lives methylomically indistinguishable (genetically determined), but grow ever more discordant throughout the life course [13]. Additionally, it has been observed that unrelated spouse-pairs that share an environment have more similar methylation profiles than those who live apart [14]. Methylomic variance attributed to the internal environment, on the other hand, is most prominent in the presence of disease, as a result of disease-related metabolic and biochemical changes [15,16]. Three key aspects of the exposure-methylation-disease framework will be discussed in the sections to follow. These are also depicted in Figure 1.

Urbanization is associated with NCD risk

The association between urbanization and NCDs is predominantly driven by characteristics of the urban environment and behavioral factors. Urbanization-associated environmental factors known to increase NCD risk include increased exposure to pollution and occupational toxins [3,4]. Urbanization-associated behavioral factors contributing to increased NCD risk include increased food availability [17], decreased nonrecreational physical activity [18], a higher portion of energy intake from fat [19], protein [20] and processed foods [21], a reduction in relative energy from carbohydrates [19] and increased adiposity [21]. While smoking and alcohol consumption are known contributors to NCD risk, their relationship with urbanization is more complex [3]. Urbanization coincides with increased purchase of commercial tobacco products and exposure to tobacco-encouraging advertizing, while second-hand smoke inhalation tends to be reduced [20,22,23]. Although urban individuals are less likely to be subject to alcohol abuse, they are more likely to consume alcohol during their lifetime [20,24]. Urbanization is also associated with increased psychosocial stress as a result of social inequity and exclusion, job insecurity and growing concerns of violence and crime [25]. These are particularly prevalent when urbanization coincides with the growth of informal settlements within the urban landscapes [5]. Access to education, on the other hand, decreases NCD risk [26].

Urbanization is associated with DNAm

The same types of exposure related to NCD risk, described above, have also been independently associated with altered DNAm. In this context, mostly noncausal associations between exposures and DNAm have been investigated, although evidence for causal associations are accumulating (Figure 1). Evidence that DNAm is causally affected by urbanization-related exposures (i.e., the exposure alters DNAm, not *vice versa*) have been published for BMI [27] and smoking [28]. Genome-wide [29] and gene-specific [27] investigations have reported adiposity-related methylation changes. These methylation signatures have been used successfully to predict the efficacy of weight-loss interventions [30]. Smoking status, cumulative smoking exposure and smoking intensity can also be determined using DNAm as biomarker [31,32]. Smoking-associated differential DNAm is only partially reversible upon cessation [31,32]. Similarly, heavy drinking can be identified using a methylation-based biomarker [33]. Methylation differences associated with alcohol consumption seem to be completely reversible, indicating possible causality [33].

Regarding noncasual associations, dietary patterns, such as high fat [34] and western diets [35], have been associated with methylation differences in genes involved in lipid metabolism, adipogenesis, inflammation and glucose regulation. Physical activity-based intervention studies reported beneficial methylation and transcription changes in genes related to longevity, inflammation and metabolism in blood, skeletal muscle and adipose tissue [36–38]. Methylation levels at specific CpGs have also been incorporated into methylation-derived biological age predictors [39,40]. The discrepancy between DNAmAge and chronological age is referred to as biological/methylation age acceleration (DNAmAgeAccel) and, when positive, is used as a biomarker of accelerated cellular aging [39,40]. Urbanicity-related factors such as adiposity [41], meat consumption [41] and cigarette smoking [42] are associated with DNAmAgeAccel. Alcohol associates with DNAmAge in a nonlinear manner, where light and heavy drinkers experience DNAmAgeAccel and moderate alcohol consumers have a relative deceleration of DNAmAge [43,44]. Education, aligned with its negative association with NCD risk [26], also protects against DNAmAgeAccel [42,43].

In terms of the urban environment, a vast amount of literature has reported on the genome-wide, global and gene-specific DNAm associations with exposure to general pollution and distinct pollutants [45]. Increased exposure to traffic-related air pollution, for example, has been associated with altered methylation at the *TET1* gene, which encodes a key enzyme involved in DNA demethylation [46]. A dose-response association between traffic-related pollution and DNAm changes has also been observed [47]. Accelerated DNAmAge has also been observed in groups exposed to pollution and pesticides [48]. From a social environment point of view, neighborhood unity, aesthetics and safety have been associated with favorable DNAm and downstream expression changes in particularly stress-

and inflammation-related genes [49,50]. Such neighborhood characteristics also enable outdoor recreational physical activity that in itself has proven beneficial [5], although these characteristics are likely to be largely present in urban-dwellers of high socioeconomic status [5,51].

DNAm is associated with NCDs

Associations between DNAm and NCDs have been reported in both directions (Figure 1), as exposure-related differential DNAm may precede disease (DNAm being on the causal pathway between exposure and disease), and, conversely, disease-related metabolic changes can affect DNAm (DNAm as a noncausal biomarker of disease). Investigations into the potential causal influence of DNAm on Type II diabetes [52] and CVD development [53] are increasing, although conclusive evidence is yet to be published. As noncausal biomarkers, both DNAm and DNAmAge have been used to identify several subtypes of cancer [54,55] and CVD [56,57]. Tumorous tissues are epigenetically older than their noncancerous counterparts [58]. As a prognostic marker, DNAm, particularly DNAmAge, has been useful in predicting cancer incidence [59,60] and survival [61], cardiac events [62], premature CVD [40,63] and all-cause mortality [60] independent of traditional risk factors. Last, as an intervention strategy, methylation-altering drugs are proving to be increasingly successful in the treatment of CVD [64], cancer [65,66] and Type II diabetes [67].

The missing link

Collectively, the evidence summarized in the previous sections highlights the potential role of DNAm as a mediator between urbanization and NCDs. The only robust evidence that has been able to link the environment to DNAm, and then DNAm to disease concerns the relationship between smoking and bladder cancer in postmenopausal women [68]. Preliminary findings have, however, associated BMI-related changes in DNAm with cardio-metabolic disease development [29]. In addition, a randomized controlled trial has provided some evidence of DNAm mediating the association between exposure to air pollution and adverse cardiovascular profiles [69].

The main gaps remaining in the literature, and the best way to address them, are the topic of interest of this review. Key questions include: Have we identified all the key risk factors in the urbanicity-NCD relationship? How does the research currently address the amalgamated risk posed by the entirety of the urban versus rural context? How can we investigate and understand the role of DNAm in this lifestyle-disease model better? Thus far, the role of DNAm in NCDs has been investigated typically by focusing on one form of exposure at a time. Investigating DNAm in the context of urbanization provides the opportunity to aggregate NCD-related exposures to provide not only a more accurate reflection of the amalgamated disease risk associated with urbanicity but also to start identifying currently unknown contributing factors that explain the variance in risk after accounting for all the known single forms of exposure. Such an investigation could also provide insight regarding the extent of potential additive risk compared with a washout effect when numerous methylation-altering exposures are clustered together. By identifying DNAm mediators involved in the relationship between urbanization and NCD, we might find modifiable targets for improving population health. The sections to follow evaluate the ability, strengths and weaknesses of different approaches to best answer key questions and elaborate on our current understanding of the role of DNAm in urbanization-related population health (Figure 1).

Contextualizing urbanization

Urbanization can be driven by the net movement of individuals from rural to urban residency in search of, among others, better education or healthcare, economic success, safety or food security. In this context, urbanization can be the result of individuals moving from a rural to an urban community within their own countries or to another country/culture entirely (migration). Alternatively, urbanization can occur as a specific region progressively urbanizes. There are three main epidemiological approaches that can be used to investigate the health-related consequences of urbanization: the migration, income-comparative and urban–rural divide approach.

Migration models

Migration studies are able to investigate the effects of environmental shifts in two ways. The first studies groups of similar ancestral and geographical origin, living in different countries, such as those who remained in the home country compared with those who moved to different locations [70]. These studies are useful in that they allow investigation of an altered environment while controlling for early-life exposure and ethnicity. A study of Japanese migrants, for example, reported a dramatic increase in CVD risk in the migrant compared with the nonmigrant group, providing evidence that the environmental shift increased CVD risk in this population [71]. A second

approach is the comparison of the migrant population with their host. In these cases, ethnic inequalities in NCD risk and outcomes can be assessed, for example, the major differences in the rate of specific CVD incidence and mortality between migrants from different countries and the same HIC host population [70].

The complexity of using migration models to investigate the association of urbanization with health stems from the numerous migration-associated confounders that are not part of the rural–urban shift. First, the migrant group themselves are not necessarily reflective of the group they originated from in that migrants are often better resourced to enable their migration than those staying behind [72]. Second, the circumstances of migration complicate the separation of the health effects related to the rural–urban shift, through the stress and impact of the migration itself [72,73]. Third, migrants often experience a vast shift in culture and climate. In addition, cultural differences in health-seeking behavior may lead to a lack of timely disease diagnosis or noncompliance in treatment, particularly in new migrants [74]. Last, migration timing may profoundly influence the relative severity of the effect of urbanization on health outcomes. According to the Developmental Origins of Health and Disease hypothesis, individuals born in an environment with limited nutritional resources (often rural settlements) are prenatally programmed to survive in these conditions. If they are then subsequently exposed to nutritional abundance, they are not metabolically equipped to manage this affluence and are predisposed to develop NCDs [75,76].

Income-comparative models

An alternative model that can be used to investigate urbanization is a global income-comparative research model where groups of differing demographic backgrounds from across the world are compared with respect to disease prevalence. The largest investigation currently implementing such a study design is the international Prospective Urban and Rural Epidemiology (PURE) study. The study includes 225,000 individuals residing in 27 low-, middle- and high-income countries [77]. The PURE study has provided many insights on global NCD risk progression and contributors. Meta-analyses by the PURE cohort include topics such as carbohydrate and fat intake [78] fruit, vegetable and legume intake [79], dietary nutrients [78], education [26], physical activity [80] and alcohol consumption [81] in relation to CVD and its related health outcomes such as blood lipid concentrations and blood pressure. The primary focus of these meta-analyses is better understanding of the relationship between country-income classification, subsequent exposure and the influence of this on CVD incidence and mortality.

Although these investigations provide vital information on the extent of the NCD crisis, particularly in LMICs, this approach is limited in a few ways. First, it is often unable to account for the genetic diversity risk of specific groups when comparing and combining multiple ethnicities. It is widely known that the risk models, ranges and cutoff created for one population are not always indicative of the same risk variance in other populations [82]. For this reason, continual attempts are being made by the WHO to recalibrate NCD risk models that are currently used in HICs to be used in LMIC population groups [83]. The numerous genome-wide association analyses that have indicated ethnic differences in genotype frequencies and their associations with intermediate phenotypes and ultimate risk indicate that although phenotypic risk assessment might be the most feasible, the gap in risk variability might only be fully addressed when also considering genetic contributors [2]. In the epigenetic context, methylation differences have also been reported among ethnic groups [11].

Second, as with migrant studies, differing geographical locations also introduce confounding by climate and diet [21,84]. Cross-cultural adaption of data collection methods is critical in these cases, as reference material developed for one population might leave many factors unstudied in the population to which it is applied, purely because of their absence in the reference group [85]. Many developing nations remain severely under-represented in genetic and epigenetic research, suggesting that the driver of observed DNAm associations might not have been identified or studied previously, resulting in potentially unquantifiable confounding when comparing these population groups [86].

Last, although socioeconomic status is associated with urbanicity, national economic status (such as the World Bank status used in most income-comparative models) does not reflect urbanicity. Developing nations, for example, are often LMICs, but generally have urban capitals, informal urban and rural settlements and rural agricultural landscapes [87].

Within-country rural–urban models

A third approach that can be used to investigate the process of progressive urbanization is to consider those who do not undergo urbanization during their lifetime, but are, instead, subject to the urban–rural divide still common in developing countries [51]. This research design can be used under the condition of having a cohort that represents

communities of a single genetic origin, part of which resides in an urban, and the other in a rural area, and does so for its entire lifespan. These individuals should have been born, and remained, on either the rural or urban side of the sociodemographic divide throughout their lifetime. A cohort of this nature will allow for the investigation of discrepant environmental exposure and health outcomes while limiting many of the confounding factors discussed in the previous sections.

Two examples of large-scale studies that can leverage this approach are the PURE [77] and the Research on Obesity and Diabetes among African Migrants (RODAM [88]) cohorts. Although spanning continents, all the countries participating in the PURE study contribute a variety of both rural and urban subcohorts [87]. The RODAM study, on the other hand, includes a rural and an urban site in Ghana (Africa), in addition to the Ghanaian migrants residing in Europe [88].

Only one of the PURE subcohorts has published epigenetic data [89], but many of the other subcohorts have access to previously collected peripheral blood samples in cryo-storage facilities [77]. No epigenetic data from the PURE cohort have been used to investigate urban–rural disparities. The advantage of a cohort such as PURE is the availability of longitudinal data on the disparity between large well-defined urban and rural communities in at least 27 countries [87]. The RODAM cohort has published genome-wide DNAm data in relation to obesity [90] and Type II diabetes [91], although no urban–rural epigenetic comparisons have been made to date. Although the RODAM study is currently of cross-sectional design, there are plans to transform it into a longitudinal cohort [92]. The PURE cohort was established in 2003, and the RODAM cohort in 2012. These cohorts are, therefore, able to capture urbanization at the pace it is currently experienced [77].

The country-specific urban–rural research platform has the benefit of being able to investigate the clusters of types of exposure that represent rural or urban living, while factors such as genetics, climate and geographical influencers remain constant and similar between groups. Developing nations, such as those included in the PURE and RODAM studies, are particularly likely to benefit from this approach, as the urban–rural divide is most severe in these countries. Furthermore, particularly in the context of epigenetic epidemiology, these countries often contain many under- or unstudied ethnic groups. Currently, most of the available evidence on NCDs, NCD risk factors and the role of epigenetics originates from study populations in developed countries [83,93]. As there are vast genomic and socioeconomic differences between these countries and the ethnic groups they contain [2,11], the feasibility of simply extrapolating findings from what is largely HIC European literature is unknown. Inclusion of more LMICs in large-scale research efforts will, therefore, not only provide an opportunity to generate population-relevant information to inform prevention, detection and treatment of NCDs in these countries but will also contribute to closing the knowledge gaps in the global literature. Findings from such investigations will provide external validation of generalizable findings, while highlighting the circumstances where population-specific research is needed.

Current challenges

One of the limitations of most epigenetic investigations is the unavailability of disease-relevant tissues. It has been well established that DNAm signatures differ among tissues, although the available evidence on environment-methylation-disease patterns is almost exclusively derived from blood-based methylation investigations [94]. Urbanicity-associated DNAm changes are, therefore, more likely to be leveraged as biomarkers of exposure or disease indicators, as the unavailability of target tissues for specific disease or outcomes limits causal inference. Mediation analyses such as Mendelian randomization could be employed to help with this, although multiple causal inference methods might be needed for triangulation of evidence [52,95]. Because, to our knowledge, population-specific genomic data are not available for many LMICs currently investigated, the addition of genetic data will be a valuable contribution.

As progressive urbanization is likely to affect both the rural and urban groups in LMICs, longitudinal measurements of DNAm will be a beneficial and informative resource. Research has shown that altered environments affect health at different rates. Adiposity, for example, seems to increase rapidly once individuals relocate to urban areas, whereas fasting insulin increases at a much more gradual pace [96]. Cross-sectional representations of urbanicity and health are therefore limited, as they capture only the factors that have an impact at the specific point in time. Should longitudinal data collection be performed, not only will there be better control of the epidemiological transition over time but this will also allow researchers to address the gap in longitudinal epigenetic research in terms of causality, reversibility and/or stability of DNAm. Standardized protocols for blood collection, handling and storage will be critical in avoiding the limitation of time-point-related batch variance.

Leveraging richly phenotyped, genetically similar, rural and urban communities with genome-wide epigenetic data and the ability to track NCD risk progression and mortality prospectively provides a unique opportunity to investigate the full environment → DNAm → NCD framework where such pathways exist, and where they do not, the value of DNAm as a biomarker for either environmental exposure or existing disease risk can be evaluated.

Conclusion

When exploring the role of DNAm in the association between urbanization and the rise in NCD prevalence, the migration, income-comparative and urban–rural study designs can be particularly useful. While each of these approaches are able to contribute to the understanding of the methylation-mediated risk attributable to urbanization from a unique perspective, they are all particularly beneficial in that they typically investigate under-represented cohorts (developing nations, migrant populations and LMICs). The integration of knowledge gained from these approaches, therefore, ultimately allow for a more rounded understanding of the role of DNAm in complex disease etiology through tackling the question from different angles, while simultaneously contributing to the currently lacking ethnic and environmental diversity of epigenetic epidemiology literature.

Future perspective

As the 21st century continues to be marked by urbanization, it is essential to improve our understanding of the molecular mechanisms driving the effect of the environmental shift on NCD prevalence and incidence. The current global landscape allows numerous approaches to be taken to investigate these mechanisms, each with its own strengths, limitations and answerable questions. In the era of big data and the continual pressure of the scientific community to promote open access and increase data availability, we expect the use of these models to significantly add to the genetic and environmental diversity captured in global epigenetic epidemiology data. Integrating the knowledge gained from the different perspectives of each of these three models will allow for a more holistic view of the different genetic and environmental origins of disease and the epigenetic mechanisms that bridge them. Ultimately, it is within the rounded understanding of methylation's role in the urbanization-NCD relationship that modifiable targets can be identified to translate research to population-based NCD prevention strategies.

Executive summary

The noncommunicable disease death toll is rising globally

- In low- and middle-income countries, this is thought to be the result of urbanization.

DNA methylation could mediate the urbanization–disease relationship

- Urbanization-related exposures associate with DNA methylation (DNAm).
- DNAm associates with noncommunicable disease (NCD) risk factors and outcomes.
- Urbanization associates with NCD risk factors and outcomes.

The following models can be used to explore this hypothesis

Migration model

- This model is particularly useful when investigating genetic predisposition and the developmental origins of health and disease. The migration process itself may, however, confound associations. Keep in mind that migrants are not always representative of their native population.

Global income comparative model

- This model can be used to report the extent of global urbanization and health disparities over time, but is susceptible to genetic, cultural and climate confounding. Comparable data collection methods and reference material is a necessity when this model is used.

Within country urban–rural models

- This model significantly reduces above-mentioned confounding and is, therefore, the most controlled setting to explore the hypothesis of DNAm mediating the effect of urbanization on NCD risk. This model is likely to be suitable only for developing nations, but can, therefore, be used when investigating un- or understudied populations.

Integrating knowledge gained from these three models is the key

- Integrating the knowledge gained from the different perspectives of each of these three models will allow for a more holistic view of the different genetic and environmental origins of disease and the epigenetic mechanisms that bridge them. Ultimately, it is within the rounded understanding of methylation's role in the urbanization–NCD relationship that modifiable targets can be identified to translate research to population-based NCD prevention strategies.

Financial & competing interests disclosure

This work was supported by the South African National Research Foundation (SFH106264 to HT Cronjé) and the Newton Fund Foundation (AMS-NAF1-Pieters to M Pieters). HR Elliott works in the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council and the University of Bristol (MC_UU_00011/5). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed

No writing assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Creative Commons Attribution 4.0 License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. World Health Organization. Noncommunicable diseases. (2018). www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases
2. Pranavchand R, Reddy B. Genomics era and complex disorders: implications of GWAS with special reference to coronary artery disease, Type II diabetes mellitus, and cancers. *J. Postgrad. Med.* 62(3), 188–198 (2016).
3. Gakidou E, Afshin A, Abajobir AA *et al.* Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet* 390(10100), 1345–1422 (2017).
4. World Health Organization. World health statistics 2018: monitoring health for the SDGs, sustainable development goals. (2018). www.who.int/gho/publications/world_health_statistics/2018/en/
5. Miranda JJ, Barrientos-Gutiérrez T, Corvalan C *et al.* Understanding the rise of cardiometabolic diseases in low- and middle-income countries. *Nat. Med.* 25(11), 1667–1679 (2019).

•• **An excellent review on the broader context of noncommunicable diseases (NCDs) rising in low- and middle-income countries including environmental and macrodrivers of NCDs and valuable thoughts on how low- and middle-income countries (as part of a global community) can respond to the mounting threat.**

6. Mcgranahan G, Satterthwaite D. *Urbanisation: Concepts and Trends.* IIED Working Paper, IIED, London, UK (2014).
7. Kim J, Samaranayake M, Pradhan S. Epigenetic mechanisms in mammals. *Cell. Mol. Life Sci.* 66(4), 596–612 (2009).
8. Gaunt TR, Shihab HA, Hemani G *et al.* Systematic identification of genetic influences on methylation across the human life course. *Genome Biol.* 17(1), 61 (2016).
9. Sharp GC, Relton CL. Epigenetics and noncommunicable diseases. *Epigenomics* 9(6), 789–791 (2017).
10. Jin Z, Liu Y. DNA methylation in human diseases. *Genes Dis.* 5(1), 1–8 (2018).
11. Galanter JM, Gignoux CR, Oh SS *et al.* Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife* 6, e20532 (2017).
12. Chitrala KN, Hernandez DG, Nalls MA *et al.* Race-specific alterations in DNA methylation among middle-aged African Americans and Whites with metabolic syndrome. *Epigenetics* 15(5), 462–482 (2020).
13. Talens RP, Christensen K, Putter H *et al.* Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell* 11(4), 694–703 (2012).
14. Li S, Wong EM, Dugué P-A *et al.* Genome-wide average DNA methylation is determined in utero. *Int. J. Epidemiol.* 47(3), 908–916 (2018).
15. Ligthart S, Marzi C, Aslibekyan S *et al.* DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biol.* 17(1), 255 (2016).
16. Wahl S, Drong A, Lehne B *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 541(7635), 81–86 (2017).
17. NCD Risk Factor Collaboration. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet* 390(10113), 2627–2642 (2017).
18. World Health Organization. Prevalence of insufficient physical activity. (2018). www.who.int/gho/ncd/risk_factors/physical_activity_text/en/
19. Dehghan M, Mente A, Zhang X *et al.* Associations of fats and carbohydrate intake with cardiovascular disease and mortality in 18 countries from five continents (PURE): a prospective cohort study. *Lancet* 390(10107), 2050–2062 (2017).

20. Department of Health (South Africa), Statistics South Africa, South African Medical Research Council, & International Coach Federation. *South Africa Demographic and Health Survey 2016.* National Department of Health, Pretoria, South Africa (2019).
21. Swinburn BA, Kraak VI, Allender S *et al.* The global syndemic of obesity, undernutrition, and climate change: the Lancet Commission report. *Lancet* 393(10173), 791–846 (2019).
22. Chow CK, Corsi DJ, Gilmore AB *et al.* Tobacco control environment: cross-sectional survey of policy implementation, social unacceptability, knowledge of tobacco health harms and relationship to quit ratio in 17 low-income, middle-income and high-income countries. *BMJ Open* 7(3), e013817 (2017).
23. Brathwaite R, Addo J, Kunst AE *et al.* Smoking prevalence differs by location of residence among Ghanaians in Africa and Europe: the RODAM study. *PLoS ONE* 12(5), e0177291 (2017).
24. Dixon MA, Chartier KG. Alcohol use patterns among urban and rural residents: demographic and social influences. *Alcohol Res.* 38(1), 69–77 (2016).
25. Pikhart H, Pikhartova J. *The Relationship Between Psychosocial Risk Factors and Health Outcomes of Chronic Diseases: A Review of the Evidence for Cancer and Cardiovascular Diseases.* WHO Regional Office for Europe, Copenhagen, Denmark (2015).
26. Jacobs DR, Kromhout D. Education, diet, and incident cardiovascular disease: ecological interactions and conclusions. *Lancet Glob. Health* 7(6), e684–e685 (2019).
27. Richmond RC, Sharp GC, Ward ME *et al.* DNA methylation and body mass index: investigating identified methylation sites at HIF3A in a causal framework. *Diabetes* 65(5), 1231–1244 (2016).
28. Li S, Wong EM, Bui M *et al.* Causal effect of smoking on DNA methylation in peripheral blood: a twin and family study. *Clin. Epigenetics* 10(1), 18 (2018).
29. Mendelson MM, Marioni RE, Joehanes R *et al.* Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: a Mendelian randomization approach. *PLoS Med.* 14(1), e1002215 (2017).
30. Perez-Cornago A, Mansego ML, Zulet MA, Martinez JA. DNA hypermethylation of the serotonin receptor type-2A gene is associated with a worse response to a weight loss intervention in subjects with metabolic syndrome. *Nutrients* 6(6), 2387–2403 (2014).
31. Joehanes R, Just A, Marioni R *et al.* Epigenetic signatures of cigarette smoking. *Circ. Cardiovasc. Genet.* 9(5), 436–447 (2016).
32. Dugué P-A, Jung C-H, Joo JE *et al.* Smoking and blood DNA methylation: an epigenome-wide association study and assessment of reversibility. *Epigenetics* 15(4), 358–368 (2019).
33. Liu C, Marioni RE, Hedman ÅK *et al.* A DNA methylation biomarker of alcohol consumption. *Mol. Psychiatr.* 23(2), 422 (2018).
34. Zhang P, Chu T, Dedousis N *et al.* DNA methylation alters transcriptional rates of differentially expressed genes and contributes to pathophysiology in mice fed a high fat diet. *Mol. Metab.* 6(4), 327–339 (2017).
35. Yokoyama AS, Dunaway K, Rutkowsky J, Rutledge JC, Milenkovic D. Chronic consumption of a western diet modifies the DNA methylation profile in the frontal cortex of mice. *Food Funct.* 9(2), 1187–1198 (2018).
36. Hibler E, Huang L, Andrade J, Spring B. Impact of a diet and activity health promotion intervention on regional patterns of DNA methylation. *Clin. Epigenetics* 11(1), 133 (2019).
37. Dimauro I, Scalabrin M, Fantini C *et al.* Resistance training and redox homeostasis: correlation with age-associated genomic changes. *Redox Biol.* 10, 34–44 (2016).
38. Grazioli E, Dimauro I, Mercatelli N *et al.* Physical activity in the prevention of human diseases: role of epigenetic modifications. *BMC Genom.* 18(8), 802 (2017).
39. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 14(10), 3156 (2013).
40. Lu AT, Quach A, Wilson JG *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* 11(2), 303 (2019).
41. Dugué P-A, Bassett JK, Joo JE *et al.* Association of DNA methylation-based biological age with health risk factors and overall and cause-specific mortality. *Am. J. Epidemiol.* 187(3), 529–538 (2017).
42. Zhao W, Ammous F, Ratliff S *et al.* Education and lifestyle factors are associated with DNA methylation clocks in older African Americans. *Int. J. Environ. Res. Public Health* 16(17), 3141 (2019).
43. Quach A, Levine ME, Tanaka T *et al.* Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging* 9(2), 419–446 (2017).
44. Rosen AD, Robertson KD, Hlady RA *et al.* DNA methylation age is accelerated in alcohol dependence. *Transl. Psychiatry* 8(1), 182 (2018).
45. Martin EM, Fry RC. Environmental influences on the epigenome: exposure-associated DNA methylation in human populations. *Annu. Rev. Public Health* 39, 309–333 (2018).
46. Somineni HK, Zhang X, Myers JMB *et al.* Ten-eleven translocation 1 (TET1) methylation is associated with childhood asthma and traffic-related air pollution. *J. Allergy Clin. Immunol.* 137(3), 797–805 (2016).
47. Ding R, Jin Y, Liu X *et al.* Dose-and time-effect responses of DNA methylation and histone H3K9 acetylation changes induced by traffic-related air pollution. *Sci. Rep.* 7, 43737 (2017).

48. Ryan J, Wrigglesworth J, Loong J, Fransquet PD, Woods RL. A systematic review and meta-analysis of environmental, lifestyle, and health factors associated with DNA methylation age. *J. Gerontol. A* 75(3), 481–494 (2019).

49. Needham BL, Smith JA, Zhao W *et al.* Life course socioeconomic status and DNA methylation in genes related to stress reactivity and inflammation: the multi-ethnic study of atherosclerosis. *Epigenetics* 10(10), 958–969 (2015).

50. Smith JA, Zhao W, Wang X *et al.* Neighborhood characteristics influence DNA methylation of genes involved in stress response and inflammation: the multi-ethnic study of atherosclerosis. *Epigenetics* 12(8), 662–673 (2017).

51. Zhang XQ. The trends, promises and challenges of urbanisation in the world. *Habitat Int.* 54, 241–252 (2016).

• **Contextualizes the economic, social and environmental impacts of urbanization, which need to be well understood before focusing on any one discipline (e.g., NCDs).**

52. Elliott HR, Shihab HA, Lockett GA *et al.* The role of DNA methylation in Type II diabetes aetiology–using genotype as a causal anchor. *Diabetes* 66(6), 1713–1722 (2017).

53. Richardson TG, Zheng J, Smith GD *et al.* Mendelian randomization analysis identifies CpG sites as putative mediators for genetic influences on cardiovascular disease risk. *Am. J. Hum. Genet.* 101(4), 590–602 (2017).

54. Vrba L, Futscher BW. A suite of DNA methylation markers that can detect most common human cancers. *Epigenetics* 13(1), 61–72 (2018).

55. Patel PG, Wessel T, Kawashima A *et al.* A three-gene DNA methylation biomarker accurately classifies early stage prostate cancer. *Prostate* 79(14), 1705–1714 (2019).

56. Muka T, Koromani F, Portilla E *et al.* The role of epigenetic modifications in cardiovascular disease: a systematic review. *Int. J. Cardiol.* 212, 174–183 (2016).

57. Ward-Caviness CK, Agha G, Chen BH *et al.* Analysis of repeated leukocyte DNA methylation assessments reveals persistent epigenetic alterations after an incident myocardial infarction. *Clin. Epigenetics* 10(1), 161 (2018).

58. Hannum G, Guinney J, Zhao L *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* 49(2), 359–367 (2013).

59. Zheng Y, Joyce BT, Colicino E *et al.* Blood epigenetic age may predict cancer incidence and mortality. *EBioMedicine* 5, 68–73 (2016).

60. Levine ME, Lu AT, Quach A *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* 10(4), 573 (2018).

61. Dugué PA, Bassett JK, Joo JE *et al.* DNA methylation-based biological aging and cancer risk and survival: pooled analysis of seven prospective studies. *Int. J. Cancer* 142(8), 1611–1619 (2018).

62. Lind L, Ingelsson E, Sundström J, Siegbahn A, Lampa E. Methylation-based estimated biological age and cardiovascular disease. *Eur. J. Clin. Invest.* 48(2), e12872 (2018).

63. Perna L, Zhang Y, Mons U, Holleczek B, Saum K-U, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin. Epigenetics* 8(1), 64 (2016).

64. Watson CJ, Horgan S, Neary R *et al.* Epigenetic therapy for the treatment of hypertension-induced cardiac hypertrophy and fibrosis. *J. Cardiovasc. Pharmacol. Ther.* 21(1), 127–137 (2016).

65. Liang G, Weisenberger DJ. DNA methylation aberrancies as a guide for surveillance and treatment of human cancers. *Epigenetics* 12(6), 416–432 (2017).

66. Miranda Furtado CL, Dos Santos Luciano MC, Silva Santos RD, Furtado GP, Moraes MO, Pessoa C. Epidrugs: targeting epigenetic marks in cancer treatment. *Epigenetics* 14(12), 1164–1176 (2019).

67. Arguelles AO, Meruvu S, Bowman JD, Choudhury M. Are epigenetic drugs for diabetes and obesity at our door step? *Drug Discov. Today* 21(3), 499–509 (2016).

68. Jordahl KM, Phipps AI, Randolph TW *et al.* Differential DNA methylation in blood as a mediator of the association between cigarette smoking and bladder cancer risk among postmenopausal women. *Epigenetics* 14(11), 1065–1073 (2019).

69. Chen R, Meng X, Zhao A *et al.* DNA hypomethylation and its mediation in the effects of fine particulate air pollution on cardiovascular biomarkers: a randomized crossover trial. *Environ. Int.* 94, 614–619 (2016).

70. Agyemang C, Van Den Born B-J. Non-communicable diseases in migrants: an expert review. *J. Travel Med.* 26(2), tay1017 (2018).

71. Marmot M, Syme SL, Kagan A, Kato H, Cohen J, Belsky J. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: prevalence of coronary and hypertensive heart disease and associated risk factors. *Am. J. Epidemiol.* 102(6), 514–525 (1975).

72. Castelli F. Drivers of migration: why do people move? *J. Travel Med.* 25(1), tay040 (2018).

73. Pavli A, Maltezou H. Health problems of newly arrived migrants and refugees in Europe. *J. Travel Med.* 24(4), tax0106 (2017).

•• **Critical review regarding the wide range of health-related complications in the rapidly growing migrant and refugee populations, while providing thoughtful insights on appropriate interventions that may benefit both the host and the migrant populations.**

74. Maneze D, Digiacomo M, Salamonson Y, Descallar J, Davidson PM. Facilitators and barriers to health-seeking behaviours among Filipino migrants: inductive analysis to inform health promotion. *BioMed. Res. Int.* 2015, 5062650629 (2015).

75. Mandy M, Nyirenda M. Developmental origins of health and disease: the relevance to developing nations. *Int. Health* 10(2), 66–70 (2018).

76. Nyirenda MJ, Byass P. Pregnancy, programming, and predisposition. *Lancet Glob. Health* 7(4), e404–e405 (2019).

77. Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *Am. Heart J.* 158(1), 1–7 (2009).

78. Mente A, Dehghan M, Rangarajan S *et al.* Association of dietary nutrients with blood lipids and blood pressure in 18 countries: a cross-sectional analysis from the PURE study. *Lancet Diabetes Endocrinol.* 5(10), 774–787 (2017).

79. Miller V, Mente A, Dehghan M *et al.* Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study. *Lancet* 390(10107), 2037–2049 (2017).

80. Attaei MW, Khatib R, Mckee M *et al.* Availability and affordability of blood pressure-lowering medicines and the effect on blood pressure control in high-income, middle-income, and low-income countries: an analysis of the PURE study data. *Lancet Public Health* 2(9), e411–e419 (2017).

81. Smyth A, Teo KK, Rangarajan S *et al.* Alcohol consumption and cardiovascular disease, cancer, injury, admission to hospital, and mortality: a prospective cohort study. *Lancet* 386(10007), 1945–1954 (2015).

82. Kruger HS, Schutte AE, Walsh CM, Kruger A, Rennie KL. Body mass index cut-points to identify cardiometabolic risk in black South Africans. *Eur. J. Nutr.* 56(1), 193–202 (2017).

83. Kaptoge S, Pennells L, De Bacquer D *et al.* World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. *Lancet Glob. Health* 7(10), e1332–e1345 (2019).

- **It highlights the importance of population-specific risk assessment and provides first steps in recalibrating risk models, in this case for cardiovascular disease, to be equally accurate in high-, middle- and low-income countries.**

84. Dwyer JT, Wiemer KL, Dary O *et al.* Fortification and health: challenges and opportunities. *Adv. Nutr.* 6(1), 124–131 (2015).

85. Macintyre UE, Venter CS, Vorster HH. A culture-sensitive quantitative food frequency questionnaire used in an African population: 1. development and reproducibility. *Public Health Nutrition* 4(1), 53–62 (2001).

86. Hobbs A, Ramsay M. Epigenetics and the burden of noncommunicable disease: a paucity of research in Africa. *Epigenomics* 7(4), 627–639 (2015).

87. Corsi DJ, Subramanian SV, Chow CK *et al.* Prospective Urban Rural Epidemiology (PURE) study: baseline characteristics of the household sample and comparative analyses with national data in 17 countries. *Am. Heart J.* 166(4), 636–646 (2013).

88. Agyemang C, Beune E, Meeks K *et al.* Rationale and cross-sectional study design of the research on obesity and Type II diabetes among African migrants: the RODAM study. *BMJ Open* 4(3), e004877 (2014).

89. Cronjé HT, Elliott HR, Nienaber-Rousseau C, Pieters M. Replication and expansion of epigenome-wide association literature in a black South African population. *Clin. Epigenetics* 12(1), 6 (2020).

90. Meeks KA, Henneman P, Venema A *et al.* An epigenome-wide association study in whole blood of measures of adiposity among Ghanaians: the RODAM study. *Clin. Epigenetics* 9(1), 103 (2017).

91. Meeks KA, Henneman P, Venema A *et al.* Epigenome-wide association study in whole blood on Type II diabetes among sub-Saharan African individuals: findings from the RODAM study. *Int. J. Epidemiol.* 48(1), 58–70 (2018).

92. Agyemang C, Beune E, Meeks K *et al.* Innovative ways of studying the effect of migration on obesity and diabetes beyond the common designs: lessons from the RODAM study. *Ann. NY Acad. Sci.* 1391(1), 54–70 (2017).

93. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* 538(7624), 161–164 (2016).

94. Zhong J, Agha G, Baccarelli AA. The role of DNA methylation in cardiovascular risk and disease: methodological aspects, study design, and data analysis for epidemiological studies. *Circ. Res.* 118(1), 119–131 (2016).

95. Battram T, Richmond RC, Baglietto L *et al.* Appraising the causal relevance of DNA methylation for risk of lung cancer. *Int. J. Epidemiol.* 48(5), 1493–1504 (2019).

96. Kinra S, Andersen E, Ben-Shlomo Y *et al.* Association between urban life-years and cardiometabolic risk: the Indian migration study. *Am. J. Epidemiol.* 174(2), 154–164 (2011).